Mapping quantitative trait loci for bean traits and ovule number in *Theobroma cacao* L.

D. Clement, A.M. Risterucci, J.C. Motamayor, J. N'Goran, and C. Lanaud

Abstract: Quantitative trait loci (QTL) mapping for bean traits and the number of ovules per ovary was carried out in cocoa (*Theobroma cacao* L.) using three test-cross progenies derived from crosses between a lower Amazon Forastero male parent (Catongo) and three female parents: one upper Amazon Forastero (IMC78) and two Trinitario (DR1 and S52). RFLP (restriction fragment length polymorphism), microsatellite, and AFLP (amplified fragment length polymorphism) markers were used for mapping. Between one and six QTL for bean traits (length, weight, and shape index) and one and four QTL for the number of ovules per ovary were detected using composite interval mapping (CIM). Individual QTL explained between 5 and 24% of the phenotypic variation. QTL clusters were identified on several chromosomes, but particularly on chromosome 4. QTL related to bean traits were detected in the same region in both Trinitario parents and in a close region in the upper Amazon Forastero parent. In reference to a previous diversity study where alleles specific to Criollo and Forastero genotypes were identified, it was possible to speculate on the putative origin (Criollo or Forastero) of favorable QTL alleles segregating in both Trinitario studied.

Key words: Theobroma cacao, bean traits, quantitative trait loci (QTL).

Résumé : La cartographie de locus à caractère quantitatif (QTL) pour des caractéristiques de fèves et pour le nombre d'ovules par ovaire a été réalisée chez le cacaoyer (*Theobroma cacao* L.) à partir de trois descendances de type test-cross provenant de croisements entre un parent mâle Forastero bas Amazonien (Catongo) et trois parents femelles: un Forastero haut Amazonien (IMC78) et deux Trinitario (DR1 et S52). Des marqueurs RFLP (Restriction Fragment Length Polymorphism), microsatellite et AFLP (Amplification Fragment Length Polymorphism) ont été utilizés pour la cartographie. Un à six QTL pour des caractères de fèves (longueur, poids et index de forme) et un à quatre QTL pour le nombre d'ovules par ovaire ont été détectés par « composite interval mapping ». Ces QTL expliquaient de 5 à 24% de la variation phénotypique. Des groupes de QTL ont été identifiés sur plusieurs chromosomes, mais plus particulièrement sur le chromosome 4. Les QTL impliqués dans les caractères de fèves ont été détectés dans la même région chez les deux Trinitario et dans une région proche chez le Forastero haut Amazonien. A partir d'une étude de diversité, où les allèles spécifiques de génotypes Criollo et Forastero ont été identifiés, il a été possible de déterminer l'origine putative (Criollo ou Forastero) des allèles favorables aux QTL en ségrégation chez les deux Trinitario étudiés.

Mots clés : Theobroma cacao, caractères de fèves, locus de caractères quantitatifs.

Introduction

Theobroma cacao L. ($2n = 20$), used in the production of chocolate, is an important crop for several tropical countries, particularly in Africa where the main cocoa-producing countries are located. *Theobroma cacao* is native to Central and South America. The species is usually classified into two genetic groups, Forastero and Criollo (Cheesman 1944; Cuatrecasas 1964). It was first domesticated by the Mayans and Aztecs (Paradis 1979) who used the Criollo cocoa type. This is one of the cocoa varieties giving the finest chocolate, and it can still be found from Mexico to Colombia and Venezuela. It has a narrow genetic base (Motamayor et al.

2002). The Forastero group is composed of numerous heterogeneous wild populations and cultivated varieties that can be found from Guyana, lower Amazonia (Brazil), and the Orinoco valley (Venezuela) to upper Amazonia (Brazil, Peru, Ecuador, and Colombia). Hybrid populations between Criollo and Forastero are usually identified as the Trinitario group. In the 18th century, hybrids between traditional Criollo, called "ancient Criollo", and introduced Forastero from the lower Amazon were produced in the Caribbean region. Owing to their higher vigor and yielding capacity, Trinitario gradually spread into pure Criollo plantations, leading to further recombination between Criollo and Trinitario.

D. Clement,[1] A.M. Risterucci, and C. Lanaud. Centre de coopération internationale en recherche pour le développement (CIRAD), TA 80/02, Avenue d'Agropolis, 34398 Montpellier CEDEX 5, France.
J.C. Motamayor. USDA, 13601 Old Cutler Road, Miami, FL 33158, U.S.A.
J. N'Goran. Centre national de la recherche agronomique (CNRA), BP 808, Divo, Côte d'Ivoire.

[1]Corresponding author (e-mail: didier.clement@cirad.fr).

Numerous diversity analyses have been carried out using morphological traits (Engels 1986; Lachenaud et al. 1999) or isozyme markers (Lanaud 1988, 1999a). More recently, molecular markers have provided a more detailed analysis of the structure of *T. cacao* (Laurent et al. 1994), revealing substantial and continuous variability in the Amazon Forastero. A recent study by Motamayor et al. (2002) has shown that the Trinitario group could have been produced from a very small number of parents. Their study showed that 50% of the molecular variability of a Trinitario population analyzed from 16 countries can be explained by the combination of a single genotype of Criollo and a single lower Amazon Forastero genotype. The development of saturated genetic linkage maps has provided a powerful tool for finding quantitative trait loci (QTL) for the main agronomic traits. The first *T. cacao* linkage map was established by Lanaud et al. (1995) and a high-density linkage map with 424 markers was established by Risterucci et al. (2000).

Large and uniform bean size are traits of interest for cocoa traders and manufacturers. Variation in bean shape, length, width, thickness, and weight are related to the genetic origin of the genotype. The Criollo group generally has large, round beans, whereas Forastero beans vary in size and are rather flat. In general, bean weight is a criterion of selection in breeding programs and varieties with beans under one gram are eliminated. Genetic studies have shown relatively high heritabilities and large additive components of genetic variance for bean weight (Lockwood and Pang 1995; Fallo and Cilas 1998). The number of ovules per ovary varies between the different genetic groups. The Trinitario group generally has fewer ovules (30–40) than the Forastero group, with the highest ovule number (50–65) being found among accessions from upper Amazon Forastero (Lachenaud 1991).

The purpose of this study was to investigate the genetic basis of bean shape, weight, and the number of ovules per ovary through QTL analysis. Three genotypes were studied: two Trinitario and one upper Amazon Forastero, each of which was included in the cocoa breeding program of Côte d'Ivoire.

Materials and methods

Genetic material

The analysis involved three populations planted at the Centre National de Recherche Agronomique (CNRA) research station at Abengourou (Côte d'Ivoire) in 1981. The parental clones were DR1, S52, and IMC78 (female parents) and Catongo (male parent) for the three crosses. Parental clones were not included in the trial. DR1 and S52 are Trinitario genotypes: DR1 was selected in Djati Roenggo Estate in Java at the beginning of the 20th century. According to Lockwood and Gyamfi (1979), DR1 appears to be a hybrid between a "Caracas" Criollo type and a Forastero, and is classified as Trinitario. S52 was selected on the island of Sao Tomé. The color and morphological characters of the pods attested to a hybrid origin between Criollo and Forastero. IMC78, a Forastero from upper Amazonia, was collected by Pound (1945) near Iquitos, Peru. Catongo is a Forastero from lower Amazonia selected in the state of Bahia, Brazil (Wood and Lass 1985). Catongo is known as

an almost completely homozygous genotype. The legitimacy of individual descendants was checked with 10 microsatellite markers. The various alleles obtained for every microsatellite in the two parents and in individual descendants were compared and approximately 5% of off-types were found in each of the three progenies. After this control, the number of individuals observed and used for QTL analysis was 124, 95, and 88 for the IMC78, DR1, and S52 progenies, respectively.

Trait evaluation

Samples were collected in the 12th year after planting and only during the main harvesting period. The position of the pod on the tree and the position of beans inside the pod are factors that affect bean size and weight (Lachenaud 1995). Samples consisted of beans taken from the middle part of 10 pods produced on the trunk. Bean weight seems to be predominantly determined by the mother tree; no significant effect has been reported so far for the cross-pollinating parent (Eskes et al. 1977; Lachenaud 1991). In this study, pollination was not controlled. For individual bean measurements, a sample of 50 peeled wet beans/tree was used (beans were stored in a freezer). Bean length, width, and thickness were measured with hand-held calipers. Bean size was calculated from these measurements. A shape index was calculated as the ratio of bean length to thickness to estimate the degree of roundness of the beans. The peeled beans were weighed. The number of ovules per ovary was estimated from 50 flowers taken from the trunk. Using a stereomicroscope, the number of ovules in one carpel of each flower was counted and multiplied by a factor of five (each flower contains five carpels).

Data analysis

The mean, standard deviation, and range of length, shape index, weight, and number of ovules were estimated for the DR1, S52, and IMC78 progenies. The Shapiro and Wilk test (Shapiro et al. 1968) was used to verify normal distribution of residuals. The Newman and Keuls test was used to compare mean values of length, shape index, weight, and number of ovules at $P = 0.05$. Phenotypic correlations between analyzed traits were estimated by the Pearson's correlation coefficient. Data analyses were carried out using STATISTICA software (StatSoft Inc., Tulsa, Okla.).

Molecular markers and linkage analysis

DNA isolation

The DNA extraction procedure described by Lanaud et al. (1995) was used. DNA was purified by ultracentrifugation in a cesium chloride – ethidium bromide gradient. The amount of DNA was estimated on an agarose gel.

Molecular markers

Genetic maps were constructed using RFLP (restriction fragment length polymorphism), microsatellite (dinucleotide repeats), and AFLP (amplification fragment length polymorphism) markers.

RFLP analyses were performed according to the procedure used for Southern blotting described by Lanaud et al. (1995). One hundred eighty-one cDNA and genomic *Pst*I mapped probes were screened on Southern blots of parental

Table 1. Characteristics of the DR1, IMC78, and S52 maps and of the reference map (UPA402 × UF676).

Progenies	No. of individuals	No. of markers	No. of linkage groups	Map length (cM)	Range of group length (cM)	Average distance between two markers (cM)
DR1 × Catongo	95	192	9 (no group 6)	653	7–114	3.4
S52 × Catongo	88	138	11 (5A; 5B)	589	4–102	4.3
IMC78 × Catongo	124	223	10	721	29–109	3.2
UPA402 × UF676	181	424	10	885	66–111	2.1

DNA, restricted with *Eco*RI, *Eco*RV, *Bgl*II, *Hin*dIII, and *Xba* restriction enzymes with 3 U DNA/μg (Gibco BRL, Carlsbad, Calif.). Twenty-nine genomic probes kindly provided by Nestlé France (Crouzillat et al. 1996), were also used and screened with *Eco*RI, *Eco*RV, and *Hin*dIII restriction enzymes. CIRAD (Centre de Coopération Internationale en Recherche pour le Developpement) probes were named cTcCIR and gTcCIR, where c and g correspond to cDNA and genomic DNA, respectively. Tc corresponds to *Theobroma cacao* and CIR to CIRAD. CCG corresponds to the probes provided by Nestlé France.

AFLP analyses were performed using the Small Genome Primer Kit AFLP analysis system II (Gibco BRL, Gaithersburg, Md.) according to the procedure described by Vos et al. (1995). The selective amplification reaction was performed with an *Mse*I primer containing three selective nucleotides (N+3) and an *Eco*RI primer containing two selective nucleotides (N+2). Sixty-four AFLP primer pair combinations were screened and 21, 22, and 27 primer pairs were used for DR1, S52, and IMC78 progenies, respectively. AFLP loci were named AF X/Y, where X is the primer pair number corresponding to a specific combination of *Mse*I and *Eco*RI primers and Y is the polymorphic band number.

An enriched simple sequence repeat (SSR) library was constructed (Lanaud et al. 1999*b*) and 22 microsatellite markers were tested according to the procedure described by Risterucci et al. (2000). Microsatellite loci were named mTcCIR.

Linkage analysis

Each of the progenies studied had a highly homozygous male parent (Catongo), and a heterozygous female parent (DR1, S52 and IMC78). We had an estimation of the heterozygosity rate for Catongo given by Crouzillat et al. (1996), which was very low (2%) and heterozygosity rates were obtained in CIRAD studies (unpublished data) of 37, 27, and 27% for DR1, S52, and IMC78, respectively. Hence, the marker segregation of each progeny only reflected the heterozygosity of the female parent. Segregation of polymorphic markers was tested with a χ^2 test to verify the Mendelian segregation ratio (1:1). Linkage analyses were performed according to a test-cross model using Joinmap version 1.4 (Stam 1993). Linkage between two locus markers was determined using an LOD score of 4. Recombination frequencies were converted into map distances with the Kosambi mapping function (Kosambi 1944). The same analysis was performed using MAPMAKER/EXP version 3.0 (Lander et al. 1987) to verify the respective positions of markers in each linkage group.

Quantitative trait locus analyses

QTL analyses were carried out from maps of DR1, IMC78, and S52 with approximately one marker every 10

cM (Clément 2001). This average distance between each marker is sufficient, considering the power of QTL detection in our study (Lincoln and Lander 1992).

QTL were detected using composite interval mapping (CIM) (Zeng 1993, 1994). CIM combines interval mapping and multiple regression. It extends the simple interval mapping (SIM) method (Lander and Botstein 1989) by fitting the most significant markers outside the interval into the model, enabling more precise and efficient mapping of QTL (Zeng 1994). CIM was run with model 6 of QTL Cartographer version 1.13 (Basten et al. 1997). A forward–backward stepwise regression was carried out to choose cofactors. A minimum of 5 to a maximum of 10 of the most significant markers was taken into account. A window size of 20 cM was chosen for the test interval. Significance thresholds were defined using the method proposed by Churchill and Doerge (1994) based on permutation by random sampling of phenotypic data testing (1000 permutations). Significance levels of $\alpha = 0.10$, 0.05, and 0.01 across the genome corresponded to LOD score thresholds of 2.15, 2.40, and 3.20, respectively. The LOD score thresholds were the mean of the value assessed for each trait. The QTL were considered significant when the LOD score was greater than 2.4. A confidence interval for each QTL was computed with a value of one LOD on both sides of the peak, as described by Lander and Botstein (1989). The QTL effect was estimated for each marker by the difference ($A–H$) between the mean of homozygous (A) versus heterozygous (H) genotype individuals obtained in the test-cross progeny.

Results

Linkage group analysis

A total of 210 RFLP probes were screened. The number of segregating markers corresponding to heterozygous loci in the DR1, S52, and IMC78 progenies was 57, 47, and 40, respectively. The percentage of AFLP fragments corresponding to heterozygous loci of DR1, S52, and IMC78 were 9.0, 5.9, and 9.5, respectively, and were identified from 21, 22, and 27 primer pairs. Twenty-two microsatellites were screened and 10, 7, and 6 heterozygous loci were mapped in DR1, S52, and IMC78, respectively.

Skewed segregation was detected for 14 loci in the DR1 progeny (7.3% of the total markers) at $P = 0.05$ and for 5 loci at $P = 0.01$. Most of these loci were located in linkage group 1 (seven loci) and 4 (eight loci). A total of four loci (2.0%) showed skewed segregation in the IMC78 progeny and three loci (1.8%) showed skewed segregation in the S52 progeny.

Microsatellite and RFLP markers were used to assign linkage groups to the chromosome number of the reference map (Risterucci et al. 2000). The relative order of these markers was preserved and only a few inversions were iden-

Table 2. Statistical analyses for bean traits and ovule number.

			Statistical analysis					Correlation coefficients[d]		
Traits	Heterozygous parents	N	Mean	NK[a]	Range	SD[b]	SW[c]	Shape index[e]	Weight	Ovule no.
Length (mm)	DR1	95	24.7	a	21.6–29.0	1.58		0.33	0.80	n.s.
	S52	88	24.0	b	21.4–26.9	1.40		0.39	0.78	0.22
	IMC78	124	25.1	a	22.0–27.9	1.29		0.37	0.86	n.s.
Shape index[e]	DR1	95	3.29	a	2.82–4.16	0.25			n.s.	n.s.
	S52	88	3.44	b	2.73–4.23	0.28			n.s.	n.s.
	IMC78	124	3.85	c	3.21–4.63	0.27			n.s.	n.s.
Weight (g)	DR1	95	1.72	a	1.19–2.63	0.32				n.s.
	S52	88	1.48	b	1.03–2.7	0.25	**			0.28
	IMC78	124	1.44	b	0.92–1.90	0.18				n.s.
Ovule no.	DR1	95	49.5	b	43.7–55.0	1.8				
	S52	88	49.7	b	44.7–53.2	1.4				
	IMC78	112	58.7	a	52.7–63.5	2.1				

[a]Newman and Keuls mean comparison test. Homogeneous groups α = 0.05.
[b]Standard deviation.
[c]Shapiro and Wilk normality test. Significant deviation from normality at $P < 0.05$ (*) and $P < 0.01$ (**).
[d]Values of the Pearson coefficient at $P < 0.05$ significance; n.s., statistically non-significant at $P = 0.05$.
[e]Length:breadth.

tified with very closely linked markers. Map characteristics of the DR1, S52, and IMC78 female parents are given in Table 2 and can be compared with those of the UPA402 × U676 reference map (Risterucci et al. 2000). The 192, 138, and 223 polymorphic markers were assigned to 9 linkage groups for DR1, 11 linkage groups for S52, and 10 linkage groups for IMC78. For DR1, only one marker (mTcCIR6) was assigned to chromosome 6. Few markers were assigned to chromosomes 10 (6.5 cM) and 8 (4 cM) of the DR1 and S52 maps. Chromosome 10 was also the smallest in the IMC78 map with a length of 33 cM. For S52, chromosome 5 was split into 2 groups, 5A and 5B, with lengths of 33 and 24 cM, respectively.

Quantitative trait analysis

The analyzed traits showed a normal distribution. Statistical analyses for bean traits and ovule number are given in Table 2. Variance analyses and mean separation tests (Newman and Keuls), carried out for these traits, showed significant differences among the three progeny sets. The results (Table 2) reflected the origin of the female parents or a female–male epistatic interaction. The shape index (length:thickness) identified the more rounded shaped beans as being from the DR1 progeny, whereas those from the IMC78 progeny were longer and thinner. The correlation between bean size and ovule number was not significant, except in the S52 progeny (Table 2). No significant correlation was found between shape index, bean traits, and number of ovules per ovary.

QTL detection

QTL obtained from CIM analyses are shown in Table 3 and illustrated in Fig. 1. A total of 8, 12, and 15 QTL were detected in DR1, S52, and IMC78, respectively. For each trait, 1 to 5 significant QTL were identified. The LOD score values were between 2.7 and 11.3 (Table 3). The smallest R^2 value of the QTL explained about 5% of the phenotypic

variation, and the most important 24% (QTL involved in ovule number in IMC78).

QTL related to bean traits

A total of 27 significant QTL related to bean traits (length, shape index, and weight) were detected using CIM in the three female parents on all chromosomes except on chromosomes 7 and 10 (Table 3; Fig. 1).

QTL for bean length were detected on chromosomes 1, 2, 4, 5, 6, and 9 and the percentage of the phenotypic variation explained by these QTL varied between 4.9 and 22.3%. The most significant QTL related to bean length were localized on the same region of chromosome 4 (Table 4; Fig. 1) and explained 17.6, 22.3, and 16.2% of the phenotypic variation of DR1, S52, and IMC78, respectively. In DR1, another QTL was detected on chromosome 4 and explained 9.3% of the phenotypic variation. These two DR1 QTL had an opposite effect. In DR1 and IMC78, a QTL for bean length was detected in the same region of chromosome 2.

QTL related to the shape index were detected on the same region of chromosome 4 (Table 4) and explained 9.0, 9.3, and 6.6% of the phenotypic variation of DR1, S52, and IMC78, respectively. Other QTL related to this trait were detected on chromosomes 1 and 8 of DR1, on chromosome 5A of S52, and on chromosome 9 of IMC78 (Fig. 1).

Two QTL related to bean weight were detected in S52 and five were detected in IMC78. For these two genotypes, the most significant QTL were detected on the same region of chromosome 4 (Table 4) and explained 16.2 and 13.6% of phenotypic variation, respectively. QTL related to bean traits, detected in the same region of chromosomes 2 and 4 in both Trinitario parents, had the same genetic effect (Table 4).

QTL related to ovule number

Between one and four QTL related to the number of ovules per ovary were detected for each female parent and were located on chromosomes 1, 2, 4, 5, and 6 (Table 3; Fig. 1). QTL were detected on chromosome 4 of DR1, S52,

Table 3. QTL for bean traits and ovule number detected by CIM.

Traits	Progenies	Chromosome	Confidence interval markers[a]	CIM		
				LOD peak	R^2	Genetic effects[b]
Length	DR1	1	AF12/30–AF58/5	3.1	8.4	0.9
		2	cTcCIR53–cTcCIR249	3.9	8.2	–0.9
		4	AF55/3–mTcCIR18	4.1	9.3	–1.0
		4	CCG1419–AF13/8	6.7	17.6	1.4
	S52	2	mTcCIR11–gTcCIR151	3.1	6.3	–0.7
		4	CCG1419–AF13/8	10.4	22.3	1.4
		6	AF12/6–cTcCIR54	2.8	4.9	0.6
		9	gTcCIR102–AF50/3	3.8	8.0	0.8
	IMC78	2	cTcCIR249–AF53/4	2.8	5.0	0.6
		3	gTcCIR107–AF55/13	4.8	8.2	0.7
		4	AF15/10–gTcCIR129	8.8	16.2	–1.1
		5	gTcCIR106–AF15/11	3.1	5.0	–0.6
		6	mTcCIR6–AF12/6	5.9	10.6	0.9
Shape index	DR1	1	AF66/4–cTcCIR19	4.6	11.9	0.18
		4	CCG1419–gTcCIR129	3.3	9.0	0.16
		8	AF56/1–cTcCIR63	2.6	8.5	0.15
	S52	4	gTcCIR129–AF13/8	3.0	9.3	0.19
		5A	gTcCIR148–cTcCIR73	5.3	17.1	0.26
	IMC78	4	cTcCIR212–gTcCIR129	2.7	6.6	–0.15
		9	AF8/23–cTcCIR207	4.6	11.0	–0.20
Weight	S52	4	CCG1419–AF13/8	4.9	16.2	0.20
		9	gTcCIR102–AF50/3	3.1	10.5	0.17
	IMC78	1	AF56/6–mTcCIR15	3.5	5.7	0.09
		2	AF9/19–AF32/1	3.9	7.1	0.10
		3	gTcCIR107–AF55/13	4.0	9.0	0.11
		4	AF15/10–gTcCIR129	7.4	13.6	–0.14
		6	AF12/6–AF15/15	3.7	8.4	0.11
No. of ovules per ovary	DR1	4	gTcCIR136–gTcCIR129	6.2	23.5	1.8
	S52	1	cTcCIR46–gTcCIR151	3.7	6.3	–0.9
		2	mTcCIR3–cTcCIR249	7.9	19.5	1.3
		2	AF12/11–mTcCIR11	10.4	22.5	–1.5
		4	AF50/8–mTcCIR18	7.8	15.0	1.1
	IMC78	4	AF15/10–gTcCIR129	11.3	24.2	–2.7
		5	cTcCIR56–AF1/12	3.0	5.5	1.0
		6	AF15/15–mTcCIR19	8.0	15.5	–1.7

[a]Includes the interval tested between two markers (see Fig. 1).
[b]The QTL effect was estimated by the difference (A–H) between the mean of the two genotypes (A and H) obtained in the test-cross progeny.

and IMC78, and explained 23.5, 15.0, and 24.2% of the variation, respectively. Only the QTL detected in DR1 and IMC78 were located in the same region of this chromosome (Table 4; Fig. 1). In S52, the two most significant QTL were located on chromosome 2 and explained 19.5 and 22.5% of the phenotypic variation (Table 3). However, these two QTL had opposite genetic effects.

QTL clusters

Some QTL related to different traits were clustered together (Fig. 1). Colocalization of QTL was found on chromosome 4. Colocation between length and weight of beans was also detected on chromosome 3 in IMC78 and on chromosome 9 in S52. QTL related to bean traits and detected in clusters were also colocalized (Tables 3 and 4). Considering the significant phenotypic correlation existing between length and weight of the beans (Table 2) the colocalization of these QTL was to be expected.

Putative origin of the favorable QTL alleles identified in both Trinitario parents

Studies on Trinitario diversity have shown that the Trinitario group resulted mainly from hybridization between two almost completely homozygous Criollo and Forastero individuals (Motamayor et al. 2002) and it is most often the case that two alleles are found at each locus of Trinitario individuals. In the two Trinitario clones studied here, DR1 and S52, a heterozygous locus can carry an allele originating from Criollo, and another allele originating from Forastero. RFLP and microsatellite markers were used to study the di-

Fig. 1. Mapping of QTL detected by CIM in DR1, S52, and IMC78. Each QTL is represented by a circle located on the LOD peak. R^2 is proportional in size to the diameter of the circle. Common markers between the different maps are underlined and in bold. Linkage groups are assigned to a chromosome number on the reference map using a common specific marker.
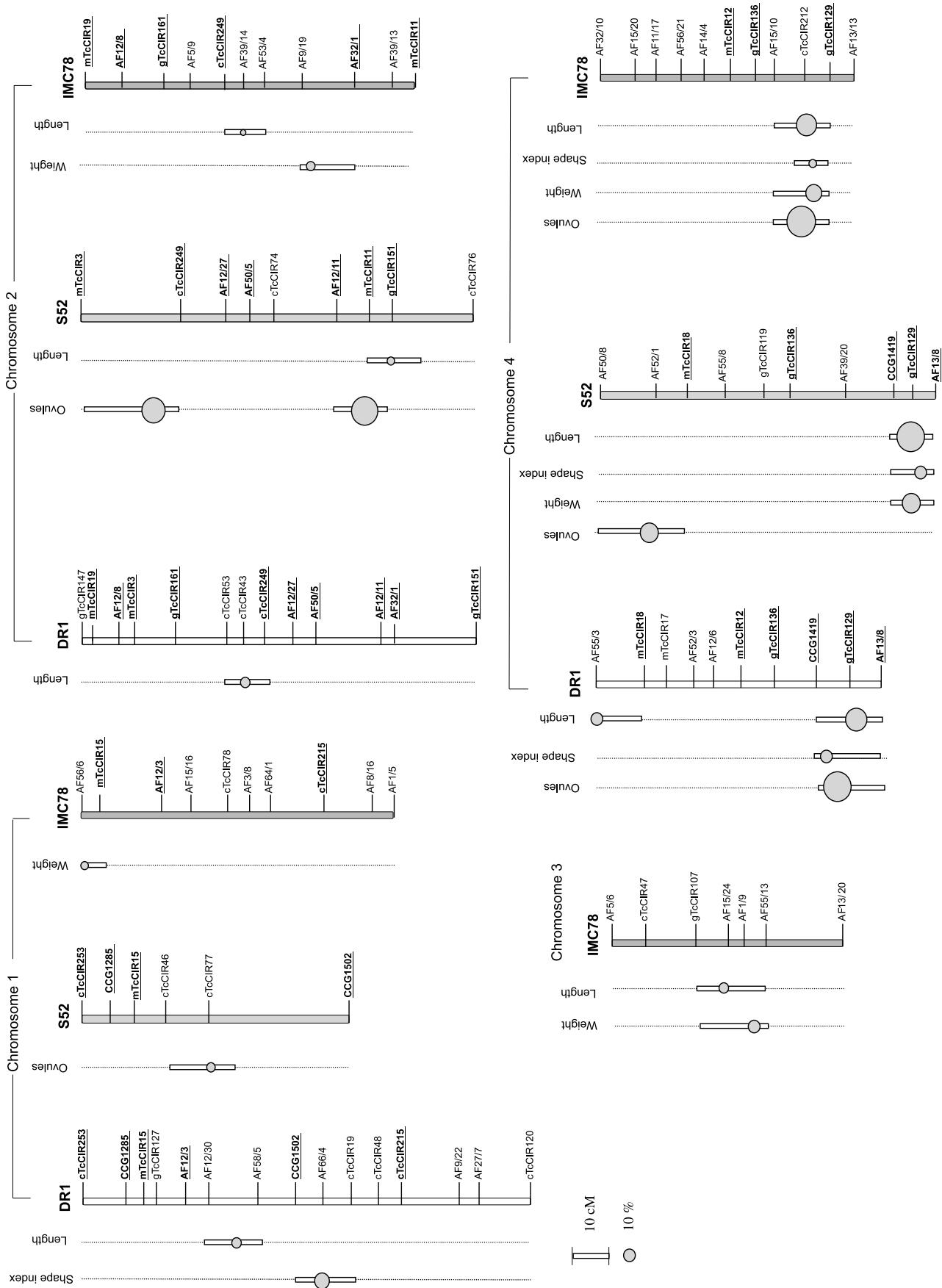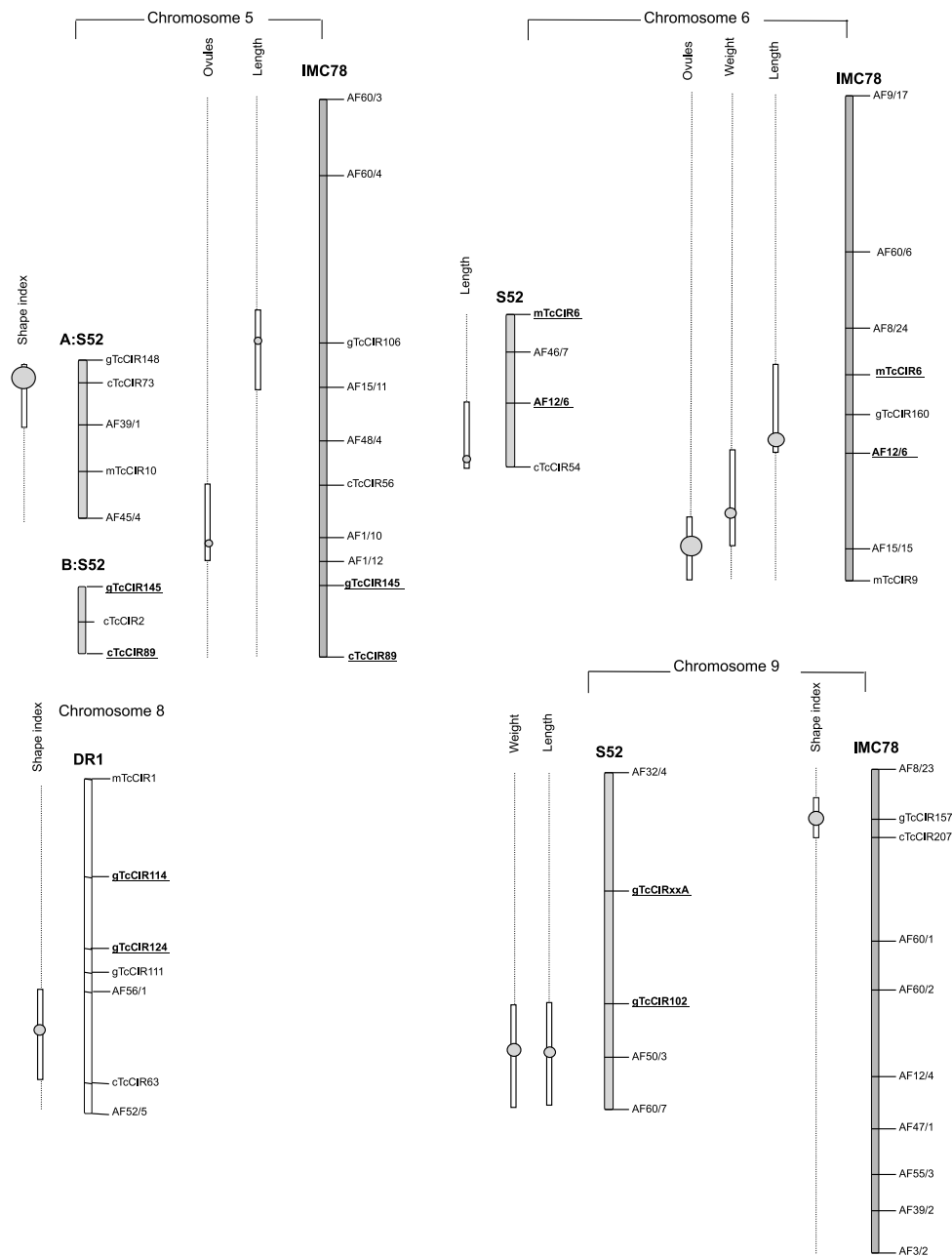
Fig. 1 (*concluded*).

versity of the Trinitario group, making it possible to distinguish the alleles specific to Criollo and Forastero, and some were mapped in DR1 and S52. For common markers used in the genetic diversity studies and in QTL analyses, the putative origin of favorable alleles at the QTL, could be determined through closely linked molecular markers (Table 5). The markers used for this analysis were chosen in the confidence interval of the QTL.

QTL related to bean length, detected on chromosomes 2 and 4 in DR1 and S52, had different effects. On chromosome 2 of DR1, the allele most probably came from Criollo. On chromosome 4 of DR1 and S52, the allele could be traced to Forastero. A similar situation was found for the alleles of the QTL related to shape index, detected on chromosomes 1 and 4 (Table 5). The shape index expresses bean

roundness. The allele of the QTL giving a shape index closer to the value of 1 had a putative Criollo origin. These results tallied with the traits of Criollo genotypes, which have round beans, whereas Forastero beans are rather long and flat, thus increasing the bean shape index values.

The favorable allele of the QTL related to ovule number, detected on chromosome 4 of DR1, had a putative Forastero origin. This result is in agreement with a generally smaller number of ovules observed in Criollo than in Forastero.

Discussion

Linkage group analysis

All types of markers were associated with skewed segregations. Skewed segregation has already been observed

Table 4. Common QTL detected on same or close regions of the genome in both Trinitario and Forastero parents.

| | | | Trinitario parents | | | | | | Forastero parent | | |
| | | | DR1 | | | S52 | | | IMC78 | | |
Traits	Marker region	Chromosome	LOD	R^2	Effect	LOD	R^2	Effect	LOD	R^2	Effect
Length	cTcCIR249	2	3.9	8.2	–0.9	—	—	—	2.8	5.0	0.6
Length	gTcCIR129	4	6.7	17.6	1.4	10.4	22.3	1.4 (A)	8.8	16.2	–1.1
Shape index	gTcCIR129	4	3.3	9.0	0.16	3.0	9.3	0.19 (A)	2.7	6.6	–0.15
Weight	gTcCIR129	4	—	—	—	4.9	16.2	0.2 (A)	7.4	13.6	–0.14
No. of ovules	gTcCIR129	4	6.2	23.5	1.8	—	—	—	11.3	24.2	–2.7

Note: Effect = (mean of the individuals with *A* genotypes) – (mean of individuals with *H* genotypes).

Table 5. Putative origin of allele in both Trinitario parents.

| | | | QTL | | | | |
| | | | DR1 | | S52 | | |
Traits	Markers	Chromosome	Origin of favorable allele	Distance to marker (cM)	Origin of favorable allele	Distance to marker (cM)	
Shape index	cTcCIR19	1	F	6			
Length	cTcCIR53–cTcCIR43–cTcCIR249	2	C	0–2–4			
Ovule no.	cTcCIR249	2			F	8	
Length	CCG1419	4	F	8	F	6	
Shape index	CCG1419	4	F	4	F	8	
Weight	CCG1419	4			F	8	
Ovule no.	CCG1419	4	F	4			
Shape index	cTcCIR63	8	F	13			

Note: F, Forastero; C, Criollo.

in other cocoa progenies (Lanaud et al. 1995; Flament 1998; Risterucci et al. 2000) with generally fewer than 10% of markers being skewed. This skewed segregation frequency was very low compared with that observed in other species like banana (36%) (Fauré et al. 1993). Gametic selection (Nakagahra 1986) or chromosomal rearrangements (Tanksley et al. 1987) have been suggested to explain distorted segregations.

The linkage groups are well distributed throughout the reference map (Risterucci et al. 2000). Several AFLP markers were located in clusters around the centromeric regions as was also observed in the reference map and in several other species such as potato (Rouppe van der Voort et al. 1997), eucalyptus (Marques et al. 1998), and tomato (Saliba-Colombani et al. 2000).

QTL detection

QTL clusters

The colocalization of QTL for bean traits and ovule number (chromosome 4 of DR1 and IMC78), could also indicate tightly linked genes, each affecting a separate character. QTL related to bean traits and ovule number explained a generally high total percentage of phenotypic variation.

Colocation with QTL identified in others studies

Some QTL identified in our study were located in the same regions as QTL identified in UF676, a Trinitario parent genotype involved in the reference genetic map. For instance, the QTL for bean weight, identified on the chromosome 9 of UF676 (N'Goran et al. 1997), is located in the

same region as the QTL identified in S52. This was also true for the QTL related to dry bean weight on chromosome 9 of UF676 (in a larger population of UPA402 × UF676) (Lanaud et al. 2000) and the QTL identified in the same region of chromosome 9 of S52. A similar situation was found for QTL related to ovule number, identified on chromosome 4 of UF676 (N'Goran et al. 1997) and on the same chromosomes in DR1 and IMC78.

Such QTL identified in the same chromosome region of several parents and in various environments could be of interest for marker-assisted selection.

Acknowledgements

References

Basten, C.J., Weir, B.S., and Zeng, Z.B. 1997. QTL cartographer: a reference manual and tutorial for QTL mapping. Ph.D. thesis, North Carolina State University, Department of Statistics, Raleigh, N.C.

Cheesman, E.E. 1944. Notes on the nomenclature, classification and possible relationships of cacao populations. Trop. Agric. **22**: 144–159.

Churchill, G.A., and Doerge, R.W. 1994. Empirical threshold values for quantitative trait mapping. Genetics, **138**: 963–971.

Crouzillat, D., Lerceteau, E., Pétiard, V., Morena, J., Rodriguez, H., Walker, D., Phillips, W., Ronning, C., Schnell, R., Osei, J., and

Fritz, P. 1996. *Theobroma cacao* L.: a genetic linkage map and quantitative trait loci analysis. Theor. Appl. Genet. **93**: 205–214.

Cuatrecasas, J. 1964. Cacao and its allies: a taxonomic revision of the genus *Theobroma.* Bull. U.S. Natl. Mus. **35**: 379–614.

Engels, J.M.M. 1986. The systematic description of cacao clones and its significance for taxonomy and plant breeding. Ph.D. thesis, Agricultural University, Wageningen, The Netherlands.

Eskes, A.B., Beek, M.A., and Toxopeus, H. 1977. Variation and correlations of some pod and bean values in cacao (*Theobroma cacao* L.) with reference to the parental effects on fruit setting and bean number. Turrialba, **27**: 193–198.

Fallo, J., and Cilas, C. 1998. Etude génétique de la granulométrie des fèves de cacaoyer (*Theobroma cacao* L.) : relation avec des caractères agronomiques. Plant. Rech. Dev. **5**: 195–200.

Fauré, S., Noyer, J.L., Horry, J.P., Bakry, F., Lanaud, C., and Gonzalez de Leon, D. 1993. A molecular marker based linkage map of diploid banana (*Musa acuminata*). Theor. Appl. Genet. **87**: 517–526.

Kosambi, D.D. 1944. The estimation of map distance from recombination values. Ann. Eugen. **12**: 172–175.

Lachenaud, P. 1991. Facteurs de la fructification chez le cacaoyer *Theobroma cacao* L. Thèse de doctorat, Ecole Nationale Supérieure Agronomique Paris–Grignon, France.

Lachenaud, P. 1995. Variations in the number of beans per pod in *Theobroma cacao* L. in the Ivory Coast. Nutritional factors, cropping effects and the role of boron. J. Hort. Sci. **70**: 7–13.

Lachenaud, P., Bonnot, F., and Oliver, G. 1999. Use of floral descriptors to study variability in wild cocoa trees (*Theobroma cacao* L.) in French Guiana. Genet. Resour. Crop Evol. **46**: 491–500.

Lanaud, C. 1988. Utilization des marqueurs enzymatiques pour l'étude génétique du cacaoyer : *Theobroma cacao* L. Contrôle génétique et "linkage" de neuf marqueurs génétiques. Café Cacao Thé, **30**: 259–270.

Lanaud, C., Risterucci, A.M., N'Goran, J., Clément, D., Flament, M., Laurent, V., and Falque, M. 1995. A genetic linkage map of *Theobroma cacao* L. Theor. Appl. Genet. **91**: 987–993.

Lanaud, C., Motamayor, J.C., and Sounigo, O. 1999*a*. Le cacaoyer. *In* Diversité génétique des plantes tropicales. *Edited by* Services des éditions du CIRAD. CIRAD, Montpellier, France. pp. 141–169.

Lanaud, C., Risterucci, A.M., Pieretti, I., Falque, M., Bouet, A., and Lagoda, P.J.L. 1999*b*. Isolation and characterization of microsatellites in *Theobroma cacao* L. Mol. Ecol. **8**: 2141–2152.

Lanaud, C., Flament, M.-H., Nyassé, S., Risterucci, A.M., Fargeas, D., Kébé, I., Motilal, L., Thévenin, J.-M., Paulin, D., Ducamp, M., Clement, D., N'Goran, J.A.K., and Cilas, C. 2000. An account of the genetic basis of cocoa tree resistance to *Phytophthora* using molecular markers. 13th International Cocoa Research Conference, Kota Kinabalu, Malaysia, 9–14 October. Cocoa Producer's Alliance, Lagos, Nigeria. In press.

Lander, E.S., and Botstein, D. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics, **121**: 185–199.

Lander, E., Green, P., Abrahamson, J., Barlow, A., Daly, M., Lincoln, S., and Newburg, L. 1987. Mapmaker: an interactive computer package for constucting primary genetic linkage maps of experimental and natural populations. Genomics, **1**: 174–181.

Laurent, V., Risterucci, A.M., and Lanaud, C. 1994. Genetic diversity in cocoa revealed by cDNA probes. Theor. Appl. Genet. **88**: 193–198.

Lincoln, S.E., and Lander, E. 1992. Systematic detection of errors in genetic linkage data. Genomics, **14**: 604–610.

Lockwood, G., and Gyamfi, M.M.O. 1979. The GRIG cocoa germplasm collection with notes on codes used in the breeding programme at Tafo and elsewhere. Technical Bulletin No. 10. Cocoa Research Institute, Tafo, Ghana.

Lockwood, G., and Pang, J.T.Y. 1995. Cocoa breeding at Bal plantations: genetic analysis and its implications for breeding strategies. First International Group for Genetic Improvement of Cocoa (INGENIC). International Workshop on Cocoa Breeding Strategies. *Edited by* INGENIC. pp. 66–80.

Marques, C.M., Araujo, J.A., Ferreira, J.G., Whetten, R., O'Malley, D.M., Liu, B.H., and Sederoff, R. 1998. AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. Theor. Appl. Genet. **96**: 727–737.

Motamayor, J.C., Risterucci, A.M., and Lanaud, C. 2002. Cacao domestication I: the origin of the cacao cultivated by the Mayas. Heredity, **88**: 380–386.

Nakagahra, M. 1986. Geographic distribution of gametophyte genes in wide crosses of rice cultivars. *In* Proceedings of the International Rice Genetics Symposium, Manila, Philippines, 27–31 May 1986. *Edited by* IRRI. pp. 73–82.

N'Goran, J.A.K., Risterucci, A.M., Clément, D., Sounigo, O., Lorieux, M., and Lanaud, C. 1997. Identification of quantitative trait loci (QTL) in *Theobroma cacao* L. Agron. Afr. **9**: 55–63.

Paradis, L. 1979. Le cacao précolombien: monnaies d'échange et breuvage des dieux. J. Agric. Trad. Bot. Appl. **26**: 3–4.

Pound, F.J. 1945. A note of the cocoa population of South America. *In* The archives of cocoa research. Vol. 1. *Edited by* H. Toxopeus. pp. 93–97.

Risterucci, A.M., Grivet, L., N'Goran, J.A.K., Pieretti, I., Flament, M.-H., and Lanaud, C. 2000. A high-density linkage map of *Theobroma cacao*. Theor. Appl. Genet. **101**: 948–955.

Rouppe van der Voort, J.N.A.M., van Zandvoort, P., van Eck, H.J., Folkertsma, R.T., Hutten, R.C.B., Draaistra, J., Gommers, F.J., Jacobsen, E., Helder, J., and Bakker, J. 1997. Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. Mol. Gen. Genet. **255**: 438–447.

Saliba-Colombani, V., Causse, M., Gervais, L., and Philouze, J. 2000. Efficiency of RFLP, RAPD, and AFLP markers for the construction of an intraspecific map of the tomato genome. Genome, **43**: 29–40.

Shapiro, S.S., Wilk, M.B., and Chen, H.J. 1968. A comparative study of various tests of normality. J. Am. Stat. Assoc. **63**: 1343–1372.

Stam, P. 1993. Construction of integrated genetic linkage maps by mean by a new computer package: JoinMap. Plant J. **3**: 739–744.

Tanksley, S.D., Miller, J., Paterson, A.H., and Bernatzky, R. 1987. Molecular mapping of plant chromosomes. *In* Chromosome structure and function. *Edited by* J.P. Gustafson and R.A. Apples. Plenum Press, New York, N.Y. pp. 157–173.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., and Zabeau, M. 1995. AFLP: a new technique for DNA. Nucleic Acids Res. **23**: 4407–4414.

Wood, G.A.R., and Lass, R.A. 1985. Cocoa. 4th ed. Longman, London, U.K.

Zeng, Z.B. 1993. Theoretical basis for separation of multiple linked gene effects in mapping quantitative trait loci. Proc. Natl. Acad. Sci. U.S.A. **90**: 10 972 – 10 976.

Zeng, Z.B. 1994. Precision mapping of quantitative trait loci. Genetics, **136**: 1457–1468.